# Running the Gamut of IoT: From Spacecraft Telemetry to Facilities Maintenance

Kyle Hundman, Data Scientist

Office of the Chief Information Officer (OCIO)

Jet Propulsion Laboratory, California Institute of Technology

# IoT, broadly

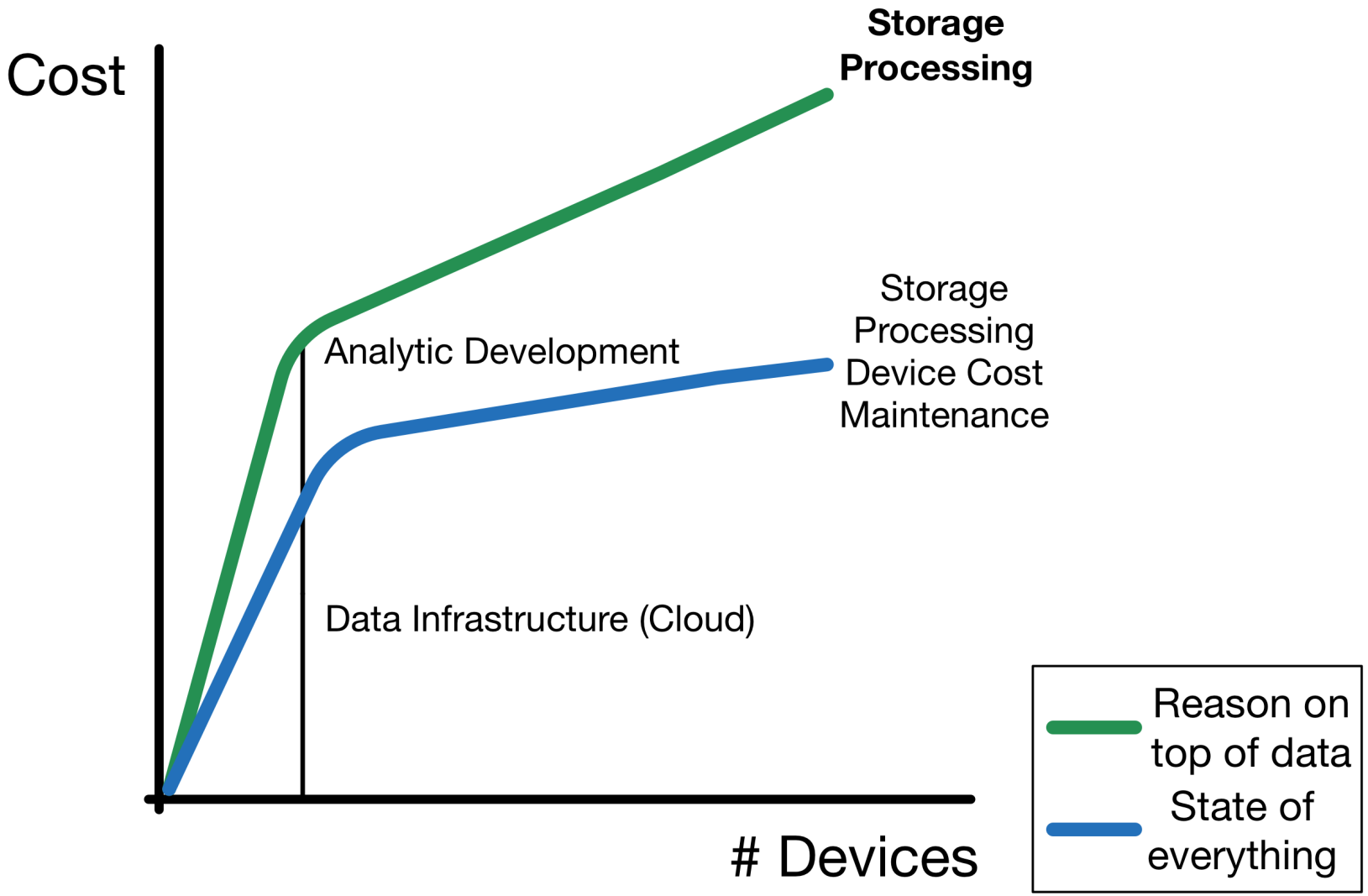
Value Approach

Dirk Didascalou, VP AWS IoT:

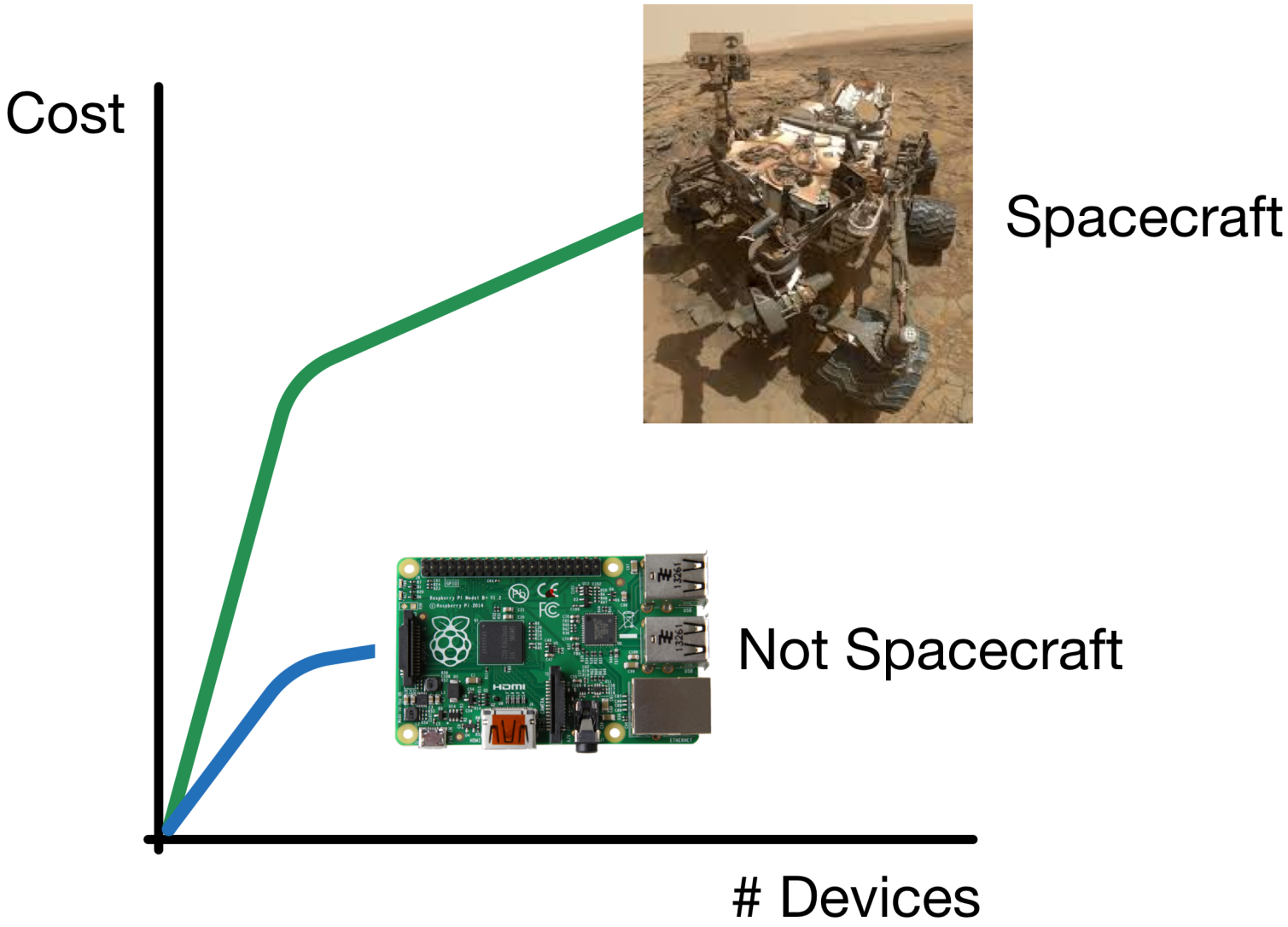
If you knew the **state of every thing** in the world, and could **reason on top of the data**:

What problems would you solve?

# Levels of Cost and Maturity



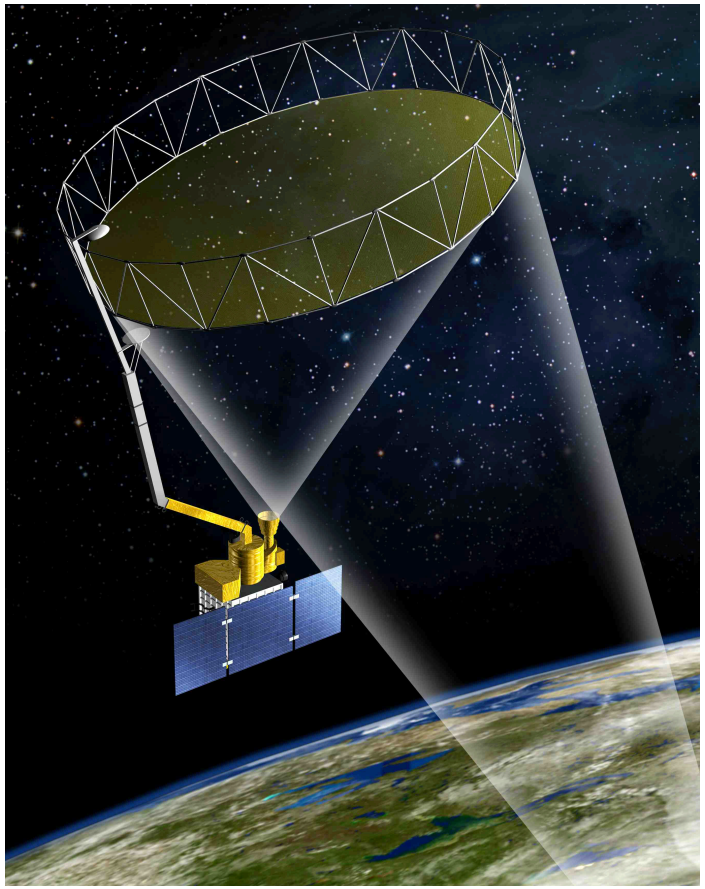
# JPL IoT: A Wide Spectrum



# Spacecraft: Megasystems of Sensors

SMAP

- Measure global soil moisture
- 3,857 Telemetry Channels
  - Power, CPU, Telecom, Radiation, Temperature, etc.
- 1 TB/day
- Near real-time processing
- Custom (limited prior knowledge)
- Complex relationships

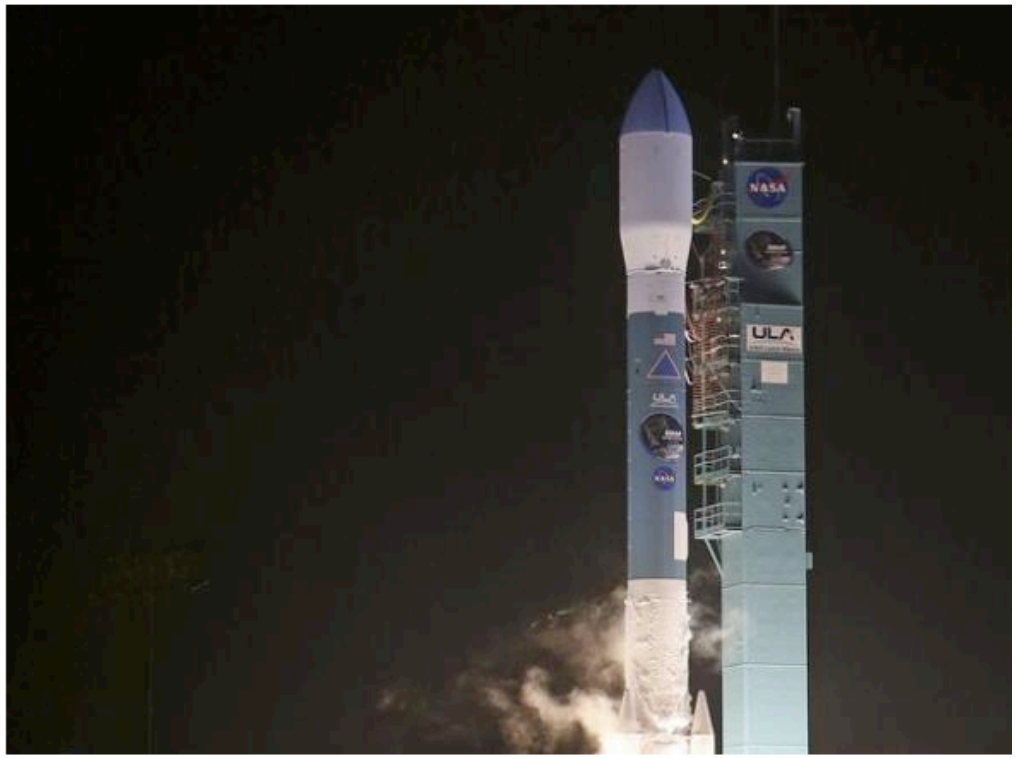


# Telemetry Anomaly Detection

## An Easy Sell

SCIENCE NEWS | Thu Sep 3, 2015 | 4:51pm EDT

### Key radar fails on $1 billion NASA environmental satellite

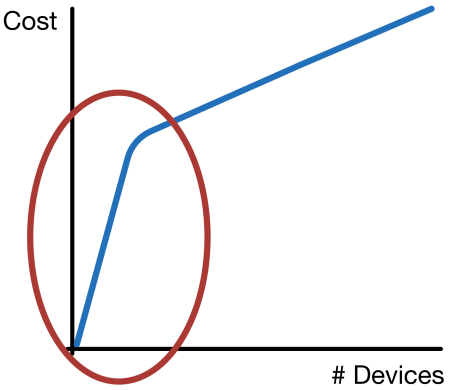


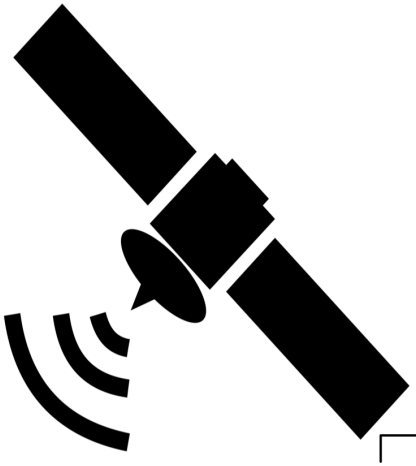

A 127-foot (39 meter) rocket built and flown by United Launch Alliance blasts off at 6:22 a.m. PST (14:22 GMT California in this January 31, 2015 file photo. REUTERS/Gene Blevins/Files

- SMAP cost: **$915M**
- Curiosity cost: **$2.5B**

- Harsh environment
- Repairs are difficult
- Public perception

# Data Infrastructure



Speed

Parallax

redis

elastic

Venue

S3

Streams

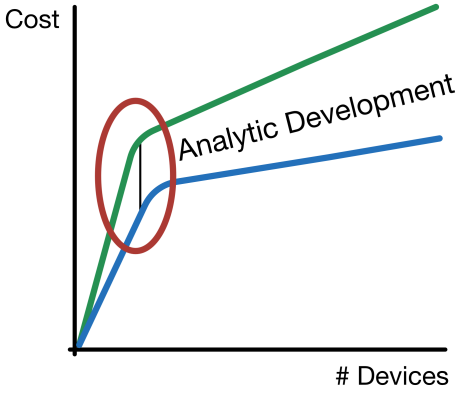
Deep Space
Network (DSN)

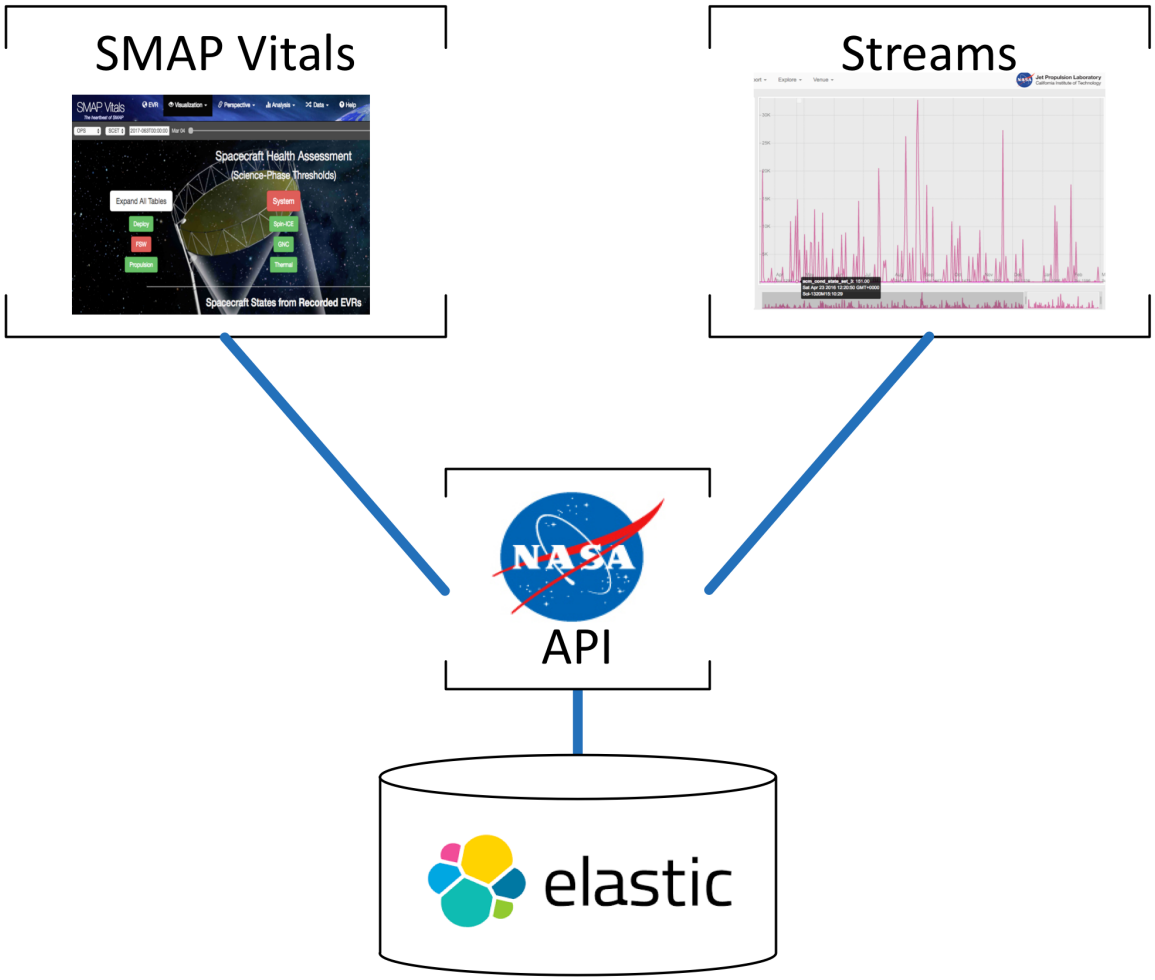
Accuracy

Credit: Dan Isla, Connor Francis

# Towards Reasoning

Flexibility, Ease of Access, Foolproofing

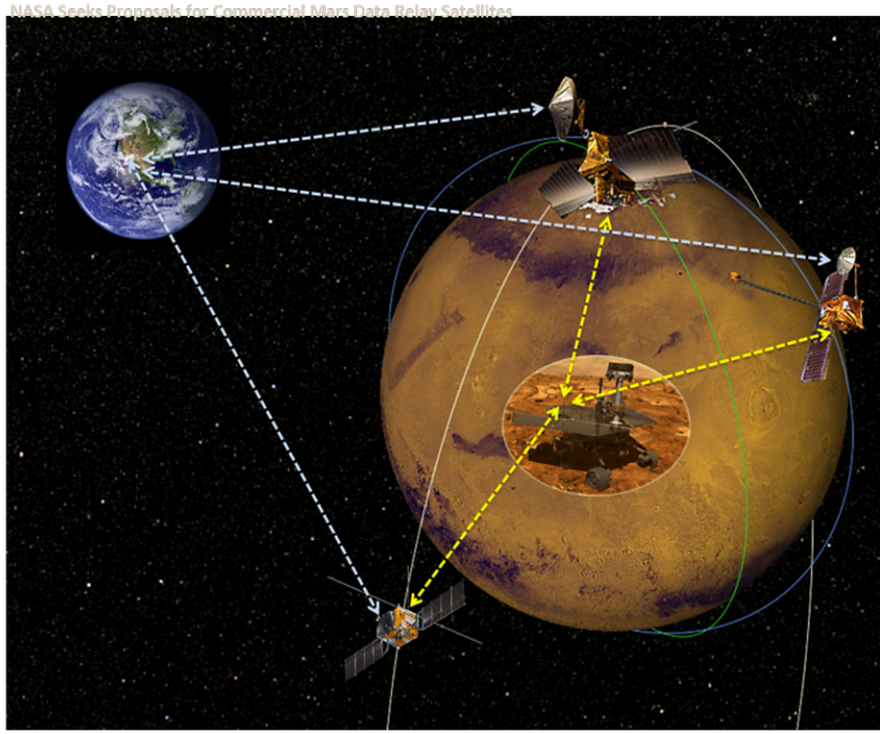


# Machine Learning Considerations

Speed and Scale

- **Speed**
  - SMAP – real-time
  - MSL (Curiosity) – 3x/day
- **Scale**
  - ~ 3.5B values
  - ~ 1M values per channel
  - Does every value matter?
    - Redundancy?
    - Importance?
    - Aggregations?

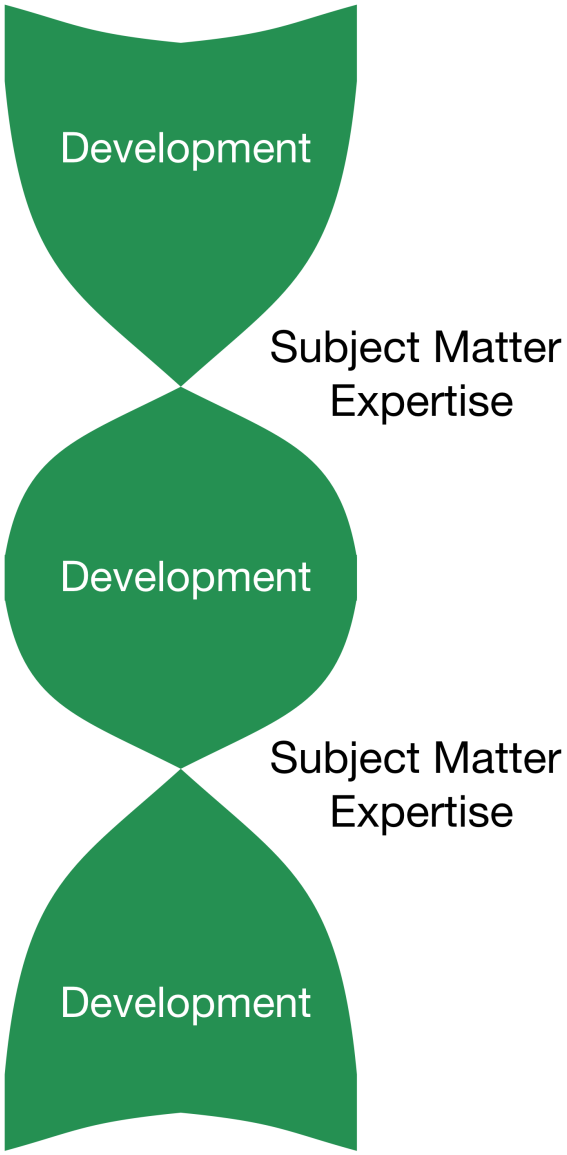


Artist rendering of commercial Mars satellites providing communications back to Earth. Image Credit: NASA/JPL

# Machine Learning Considerations

Expert Systems?

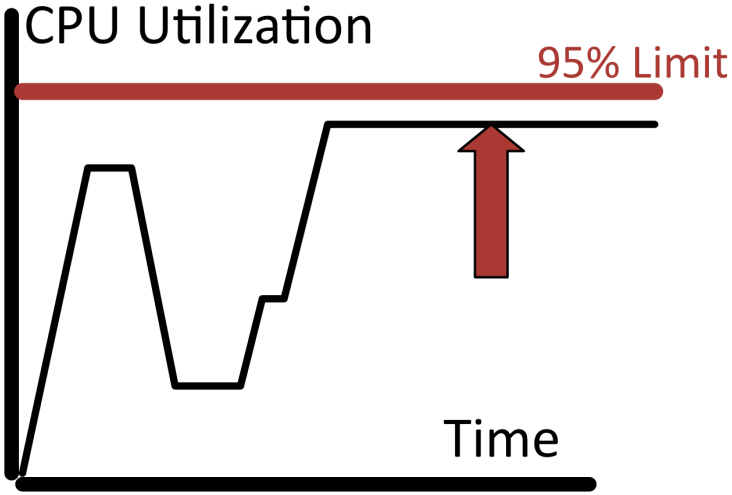
- Leverage subject matter experts
  - Parameters, limits, feature creation
- Challenges
  - Accuracy
  - Completeness
  - Time
  - Custom
- Validation, not development



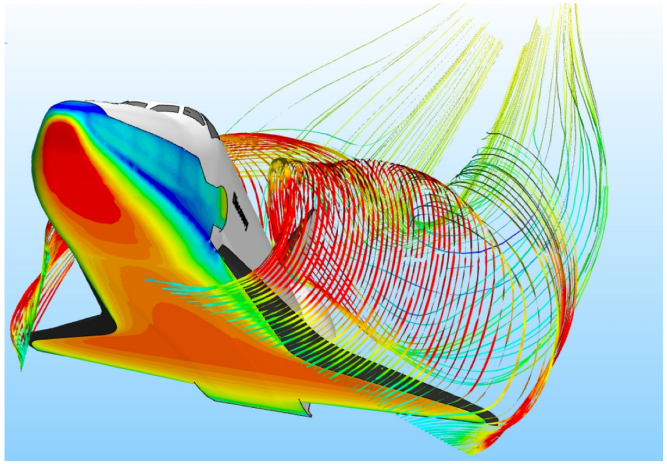
# Machine Learning Considerations

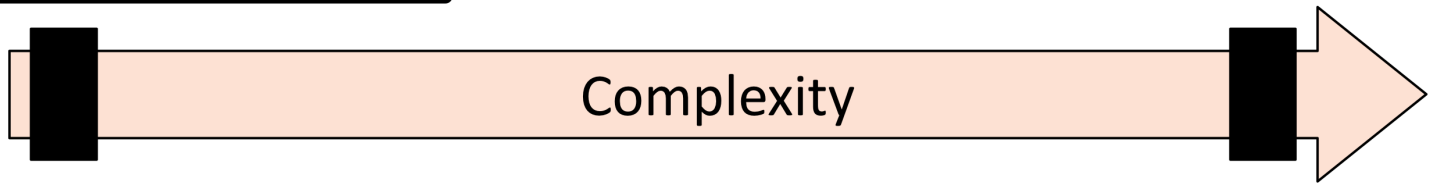
Complexity

**Limit-Checking**



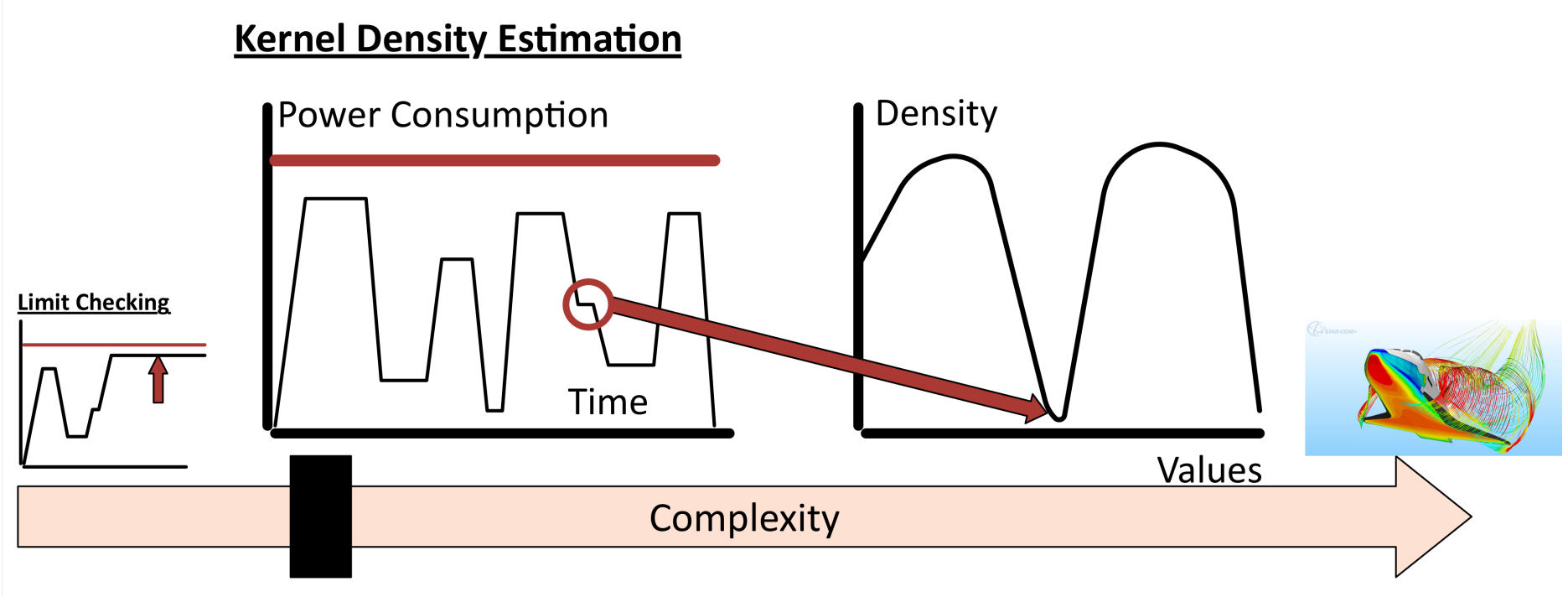

**Model-Based Simulation**





# Machine Learning Considerations

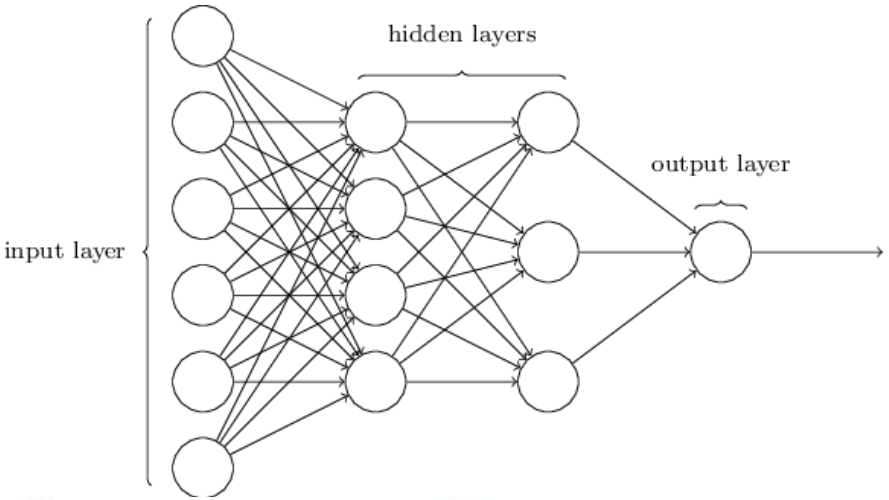
Complexity

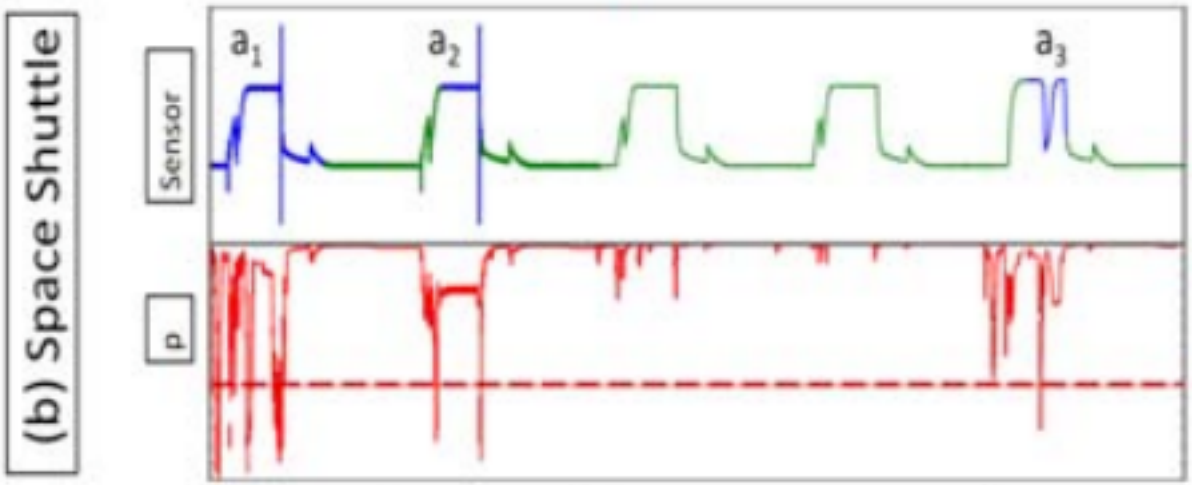


# Machine Learning Considerations

Performance and Interpretability

## Deep Learning - LSTMs



# Machine Learning Considerations

Model comparisons based on data size

Dataset Size

Model Complexity

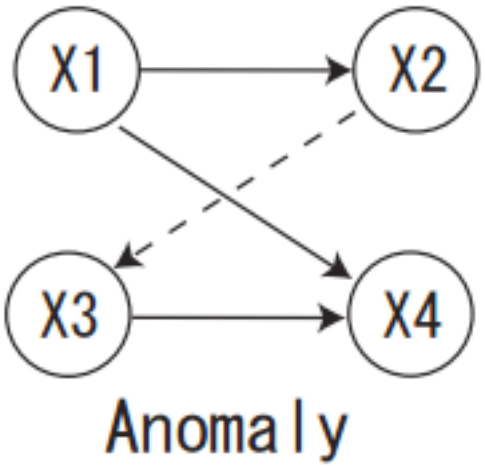
| Model | AG | Sogou | DBP. | Yelp P. | Yelp F. | Yah. A. | Amz. F. | Amz. P. |
|---|---|---|---|---|---|---|---|---|
| BoW | 11.19 | 7.15 | 3.39 | 7.76 | 42.01 | 31.11 | 45.36 | 9.60 |
| BoW TFIDF | 10.36 | 6.55 | 2.63 | 6.34 | 40.14 | 28.96 | 44.74 | 9.00 |
| ngrams | 7.96 | 2.92 | 1.37 | **4.36** | 43.74 | 31.53 | 45.73 | 7.98 |
| ngrams TFIDF | **7.64** | **2.81** | **1.31** | 4.56 | 45.20 | 31.49 | 47.56 | 8.46 |
| Bag-of-means | **16.91** | **10.79** | **9.55** | **12.67** | **47.46** | **39.45** | **55.87** | **18.39** |
| LSTM | 13.94 | 4.82 | 1.45 | 5.26 | 41.83 | 29.16 | 40.57 | 6.10 |
| Lg. w2v Conv. | 9.92 | 4.39 | 1.42 | 4.60 | 40.16 | 31.97 | 44.40 | 5.88 |
| Sm. w2v Conv. | 11.35 | 4.54 | 1.71 | 5.56 | 42.13 | 31.50 | 42.59 | 6.00 |
| Lg. w2v Conv. Th. | 9.91 | - | 1.37 | 4.63 | 39.58 | 31.23 | 43.75 | 5.80 |
| Sm. w2v Conv. Th. | 10.88 | - | 1.53 | 5.36 | 41.09 | 29.86 | 42.50 | 5.63 |
| Lg. Lk. Conv. | 8.55 | 4.95 | 1.72 | 4.89 | 40.52 | 29.06 | 45.95 | 5.84 |
| Sm. Lk. Conv. | 10.87 | 4.93 | 1.85 | 5.54 | 41.41 | 30.02 | 43.66 | 5.85 |
| Lg. Lk. Conv. Th. | 8.93 | - | 1.58 | 5.03 | 40.52 | 28.84 | 42.39 | 5.52 |
| Sm. Lk. Conv. Th. | 9.12 | - | 1.77 | 5.37 | 41.17 | 28.92 | 43.19 | 5.51 |
| Lg. Full Conv. | 9.85 | 8.80 | 1.66 | 5.25 | 38.40 | 29.90 | 40.89 | 5.78 |
| Sm. Full Conv. | 11.59 | 8.95 | 1.89 | 5.67 | 38.82 | 30.01 | 40.88 | 5.78 |
| Lg. Full Conv. Th. | 9.51 | - | 1.55 | 4.88 | 38.04 | 29.58 | 40.54 | 5.51 |
| Sm. Full Conv. Th. | 10.89 | - | 1.69 | 5.42 | **37.95** | 29.90 | 40.53 | 5.66 |
| Lg. Conv. | 12.82 | 4.88 | 1.73 | 5.89 | 39.62 | 29.55 | 41.31 | 5.51 |
| Sm. Conv. | 15.65 | 8.65 | 1.98 | 6.53 | 40.84 | 29.84 | 40.53 | 5.50 |
| Lg. Conv. Th. | 13.39 | - | 1.60 | 5.82 | 39.30 | **28.80** | 40.45 | **4.93** |
| Sm. Conv. Th. | 14.80 | - | 1.85 | 6.49 | 40.16 | 29.84 | **40.43** | 5.67 |

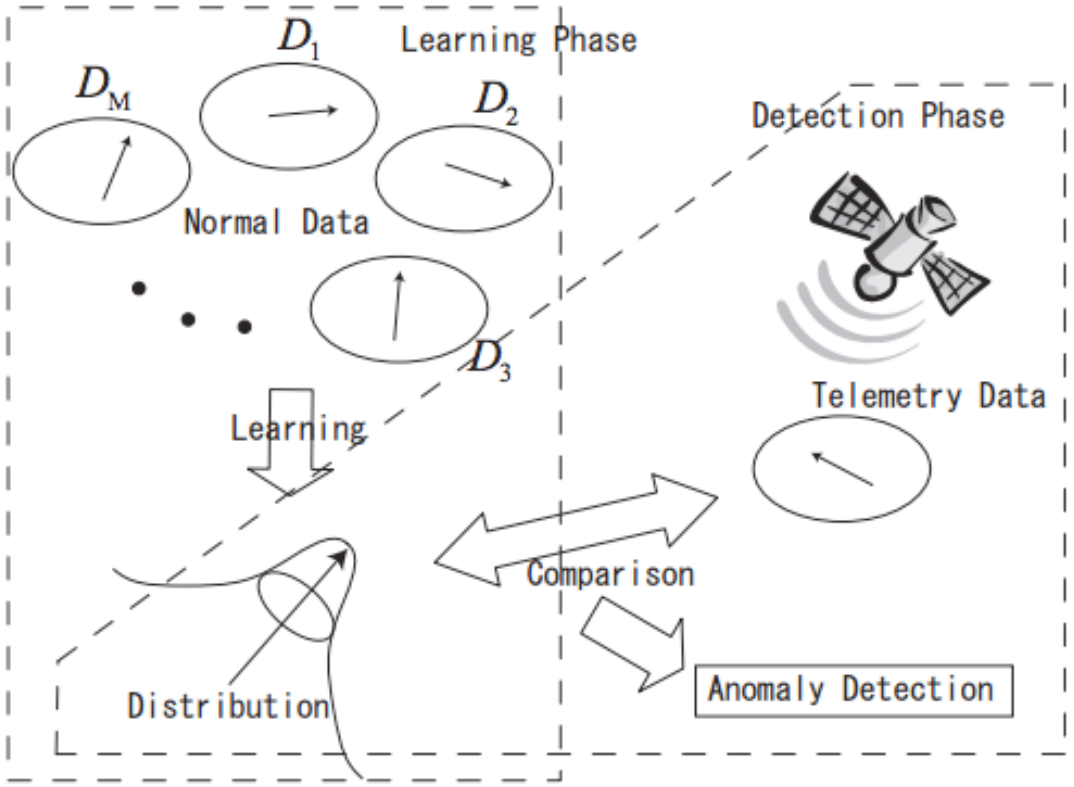
(Zhang, Zhao, LeCun 2015)

# Machine Learning Considerations

Performance and Interpretability

## Kernel PCA




(Fujimaki, Yairi, and Machida, 2005)

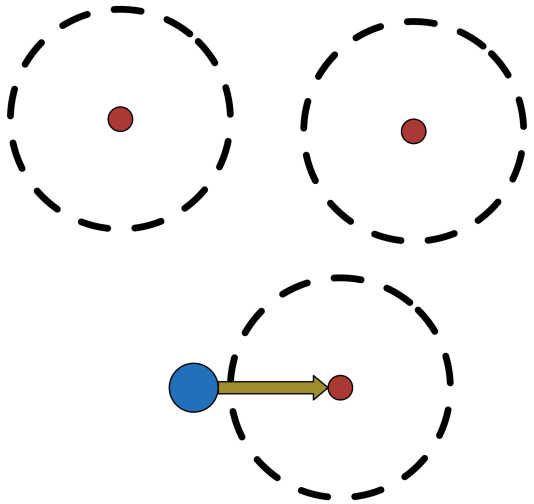
Complexity

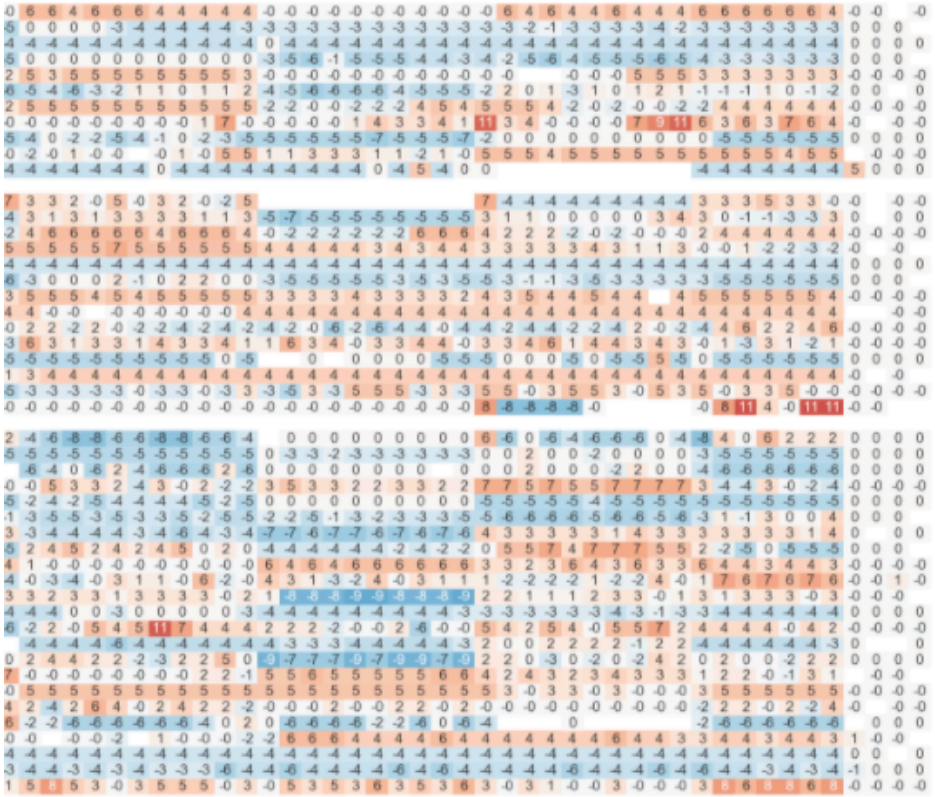
# Machine Learning Considerations

Performance and Interpretability

## K-Means Clustering

- Assumptions of normality?
- Limited interpretability
- Less expensive
- Decisions
  - Number of clusters
  - Distance threshold

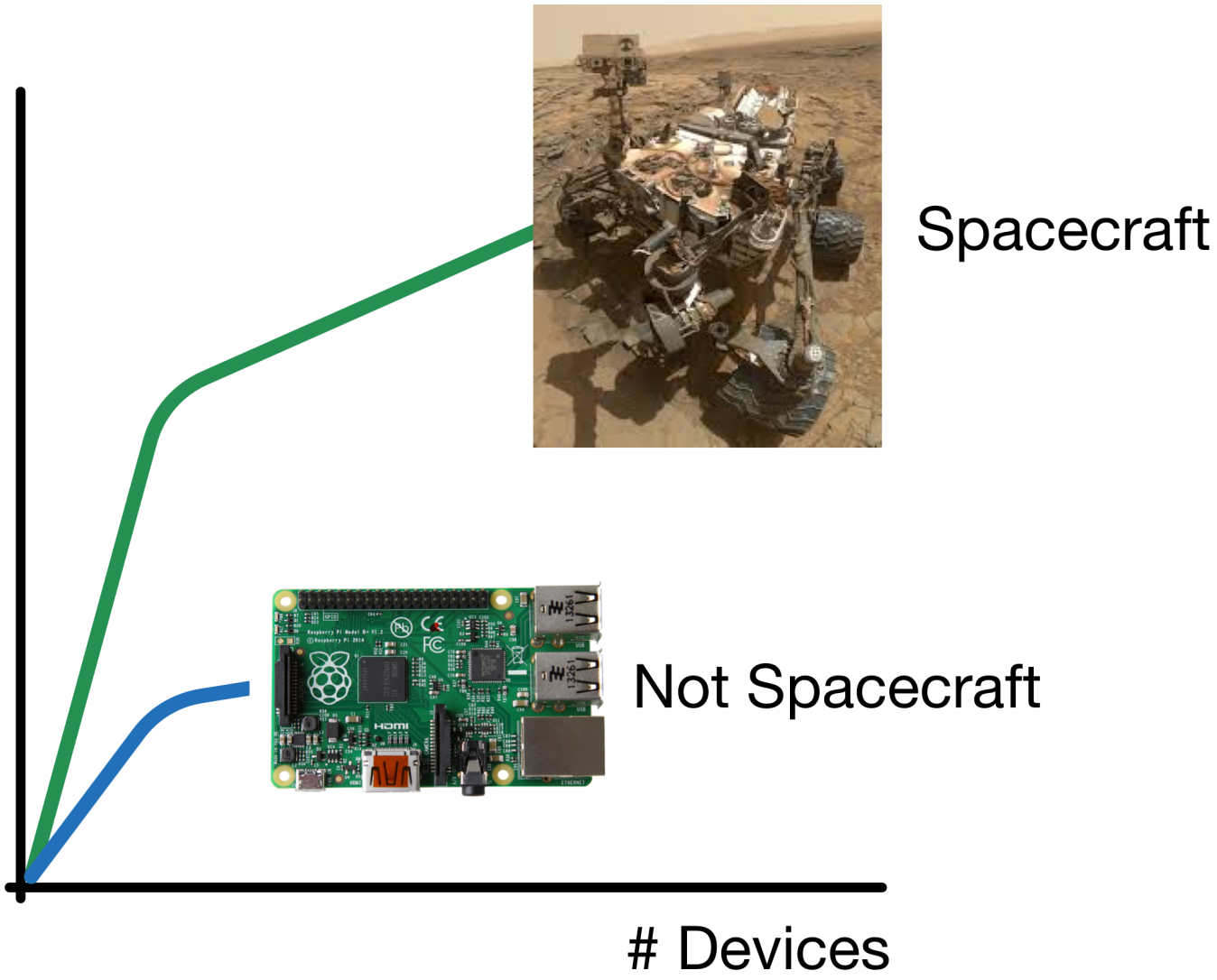




## Achieving Balance – Scientific Rigor versus Value

- Be patient and thorough
- Communication
- Good enough
- Give and take - add value along the way

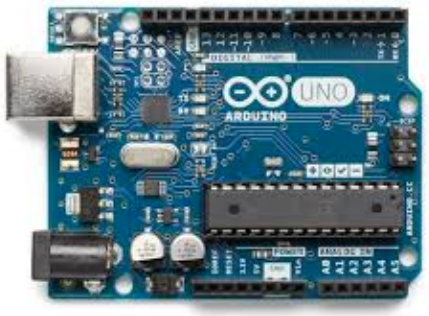
# JPL IoT:



# Leveraging the Consumer Space

Creative Possibilities

Arduino




docker

GitHub

Amazon Echo


Raspberry Pi


AWS IoT

Phillips Hue

# Makoto Koike, IoT Hero

## Cucumber Sorting

"Makoto started helping out at his parents' cucumber farm, and was amazed by the amount of work it takes to sort cucumbers by size, shape, color, and other attributes."

*Makoto Koike, center, with his parents at the family cucumber farm*

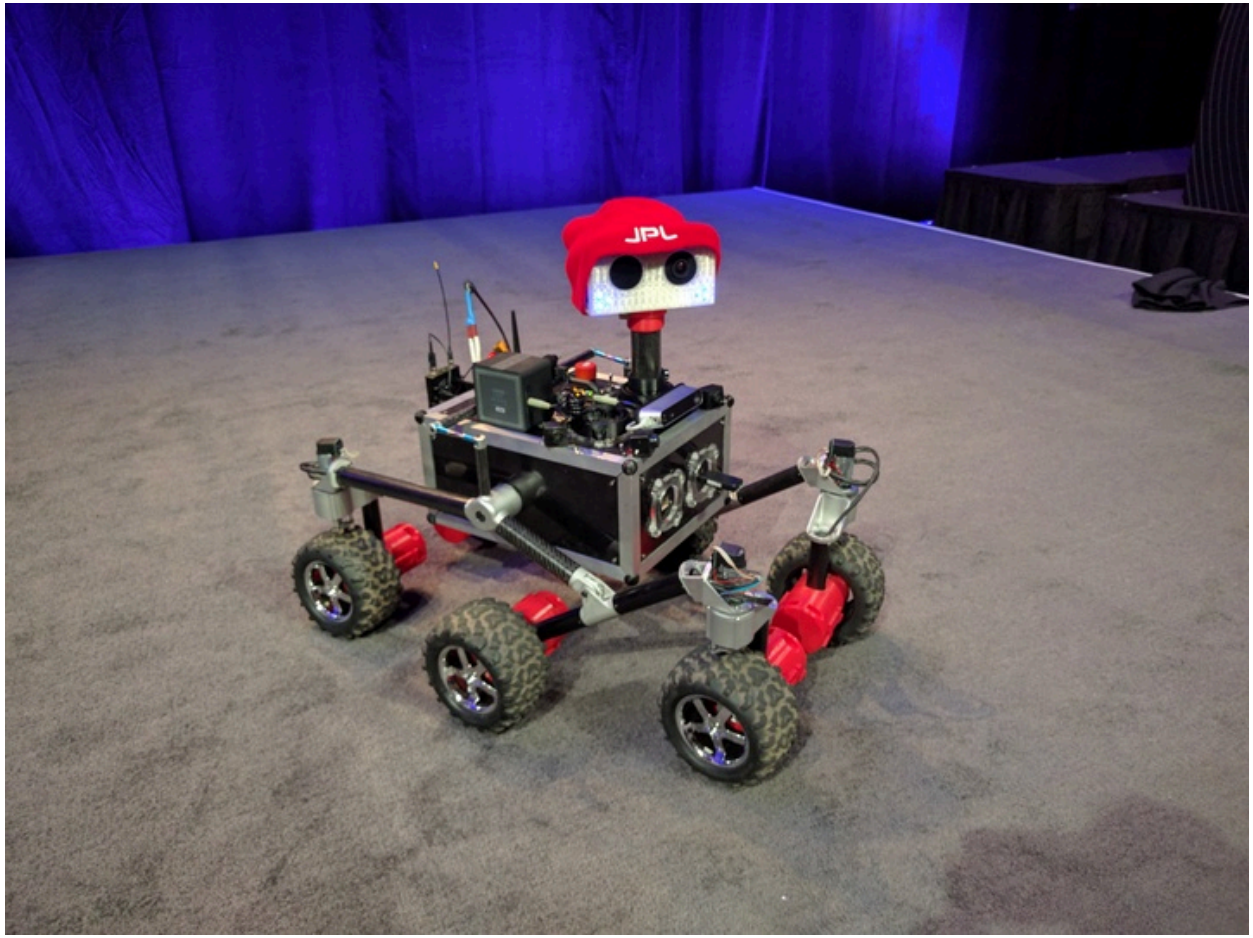
https://cloud.google.com/blog/big-data/2016/08/how-a-japanese-cucumber-farmer-is-using-deep-learning-and-tensorflow

# Cucumber Sorter


https://cloud.google.com/blog/big-data/2016/08/how-a-japanese-cucumber-farmer-is-using-deep-learning-and-tensorflow

# Rov-E

Robotic Ambassador




https://youtu.be/QcqSHLw4hTE?t=46m40s

# Space Problems

Maximizing Utilization

Motion Sensor - $4

WiFi Adapter - $9

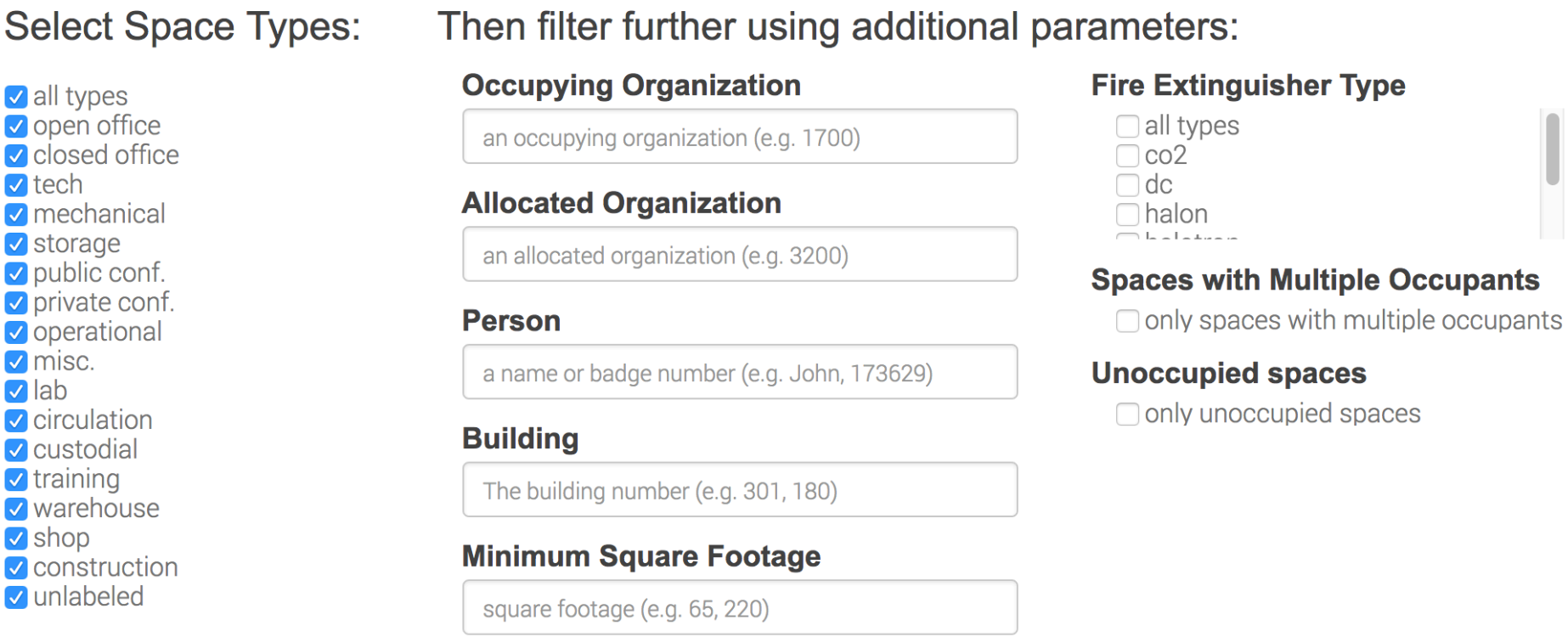

Raspberry Pi 2 - $30

DynamoDB
25GB, 2.5M reads free

DynamoDB

# FacSearch

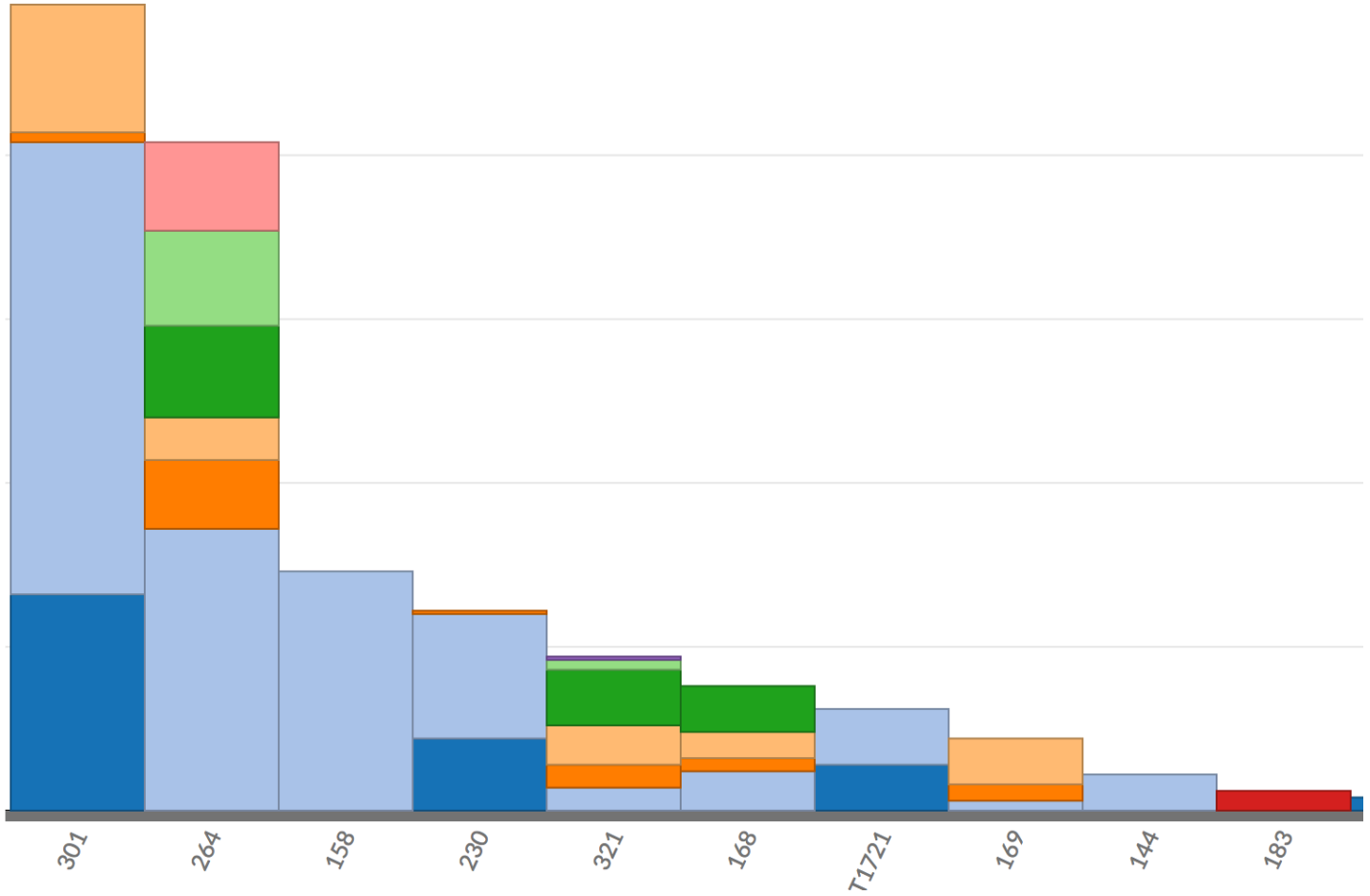
Visualization and Integration Platform for Facilities and IoT Data




https://github.com/khundman/FacSearch

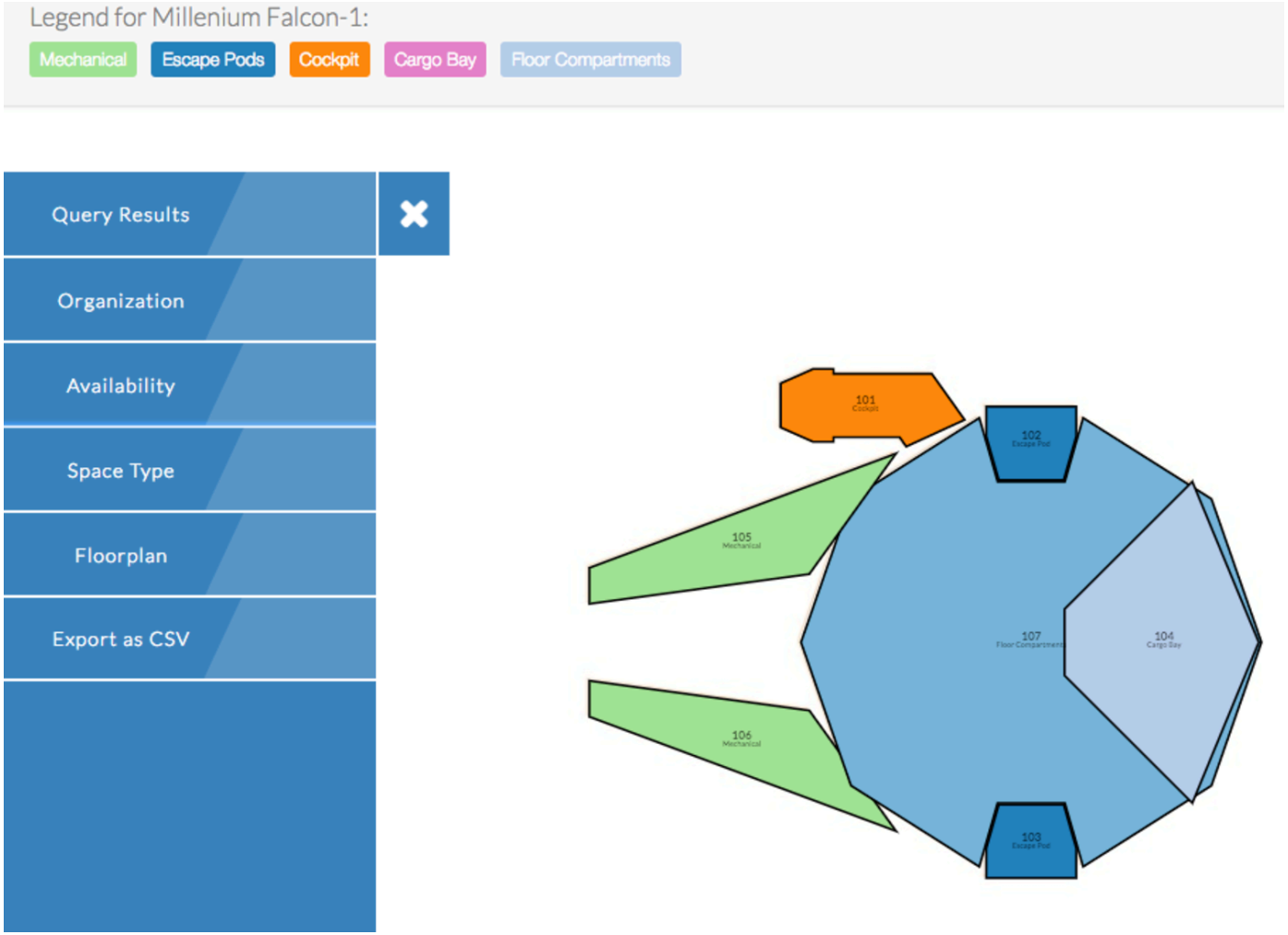
# FacSearch

Data Visualization and Integration Platform for Facilities



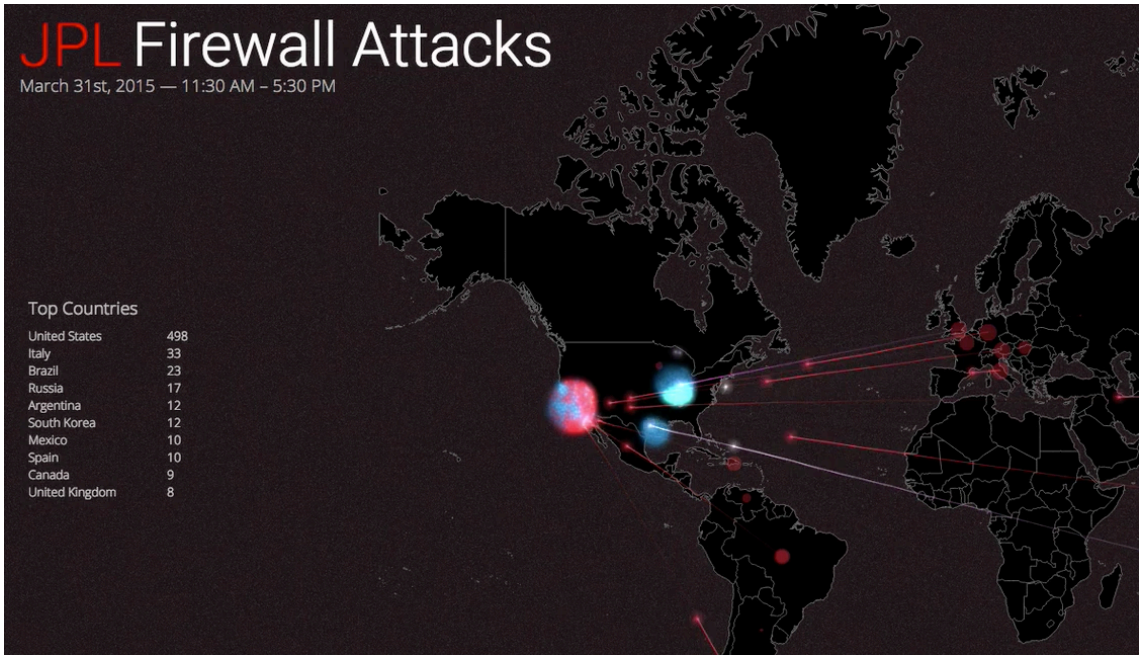
# FacSearch

Data Visualization and Integration Platform for Facilities



# Extend Situational Awareness

- Firewall attacks
- Acronyms
- Remedy tickets
- Safety
- Rov-E
- Mars Alexa




Newsletter signup

Space & Science

## NASA unveils a skill for Amazon's Alexa that lets you ask questions about Mars

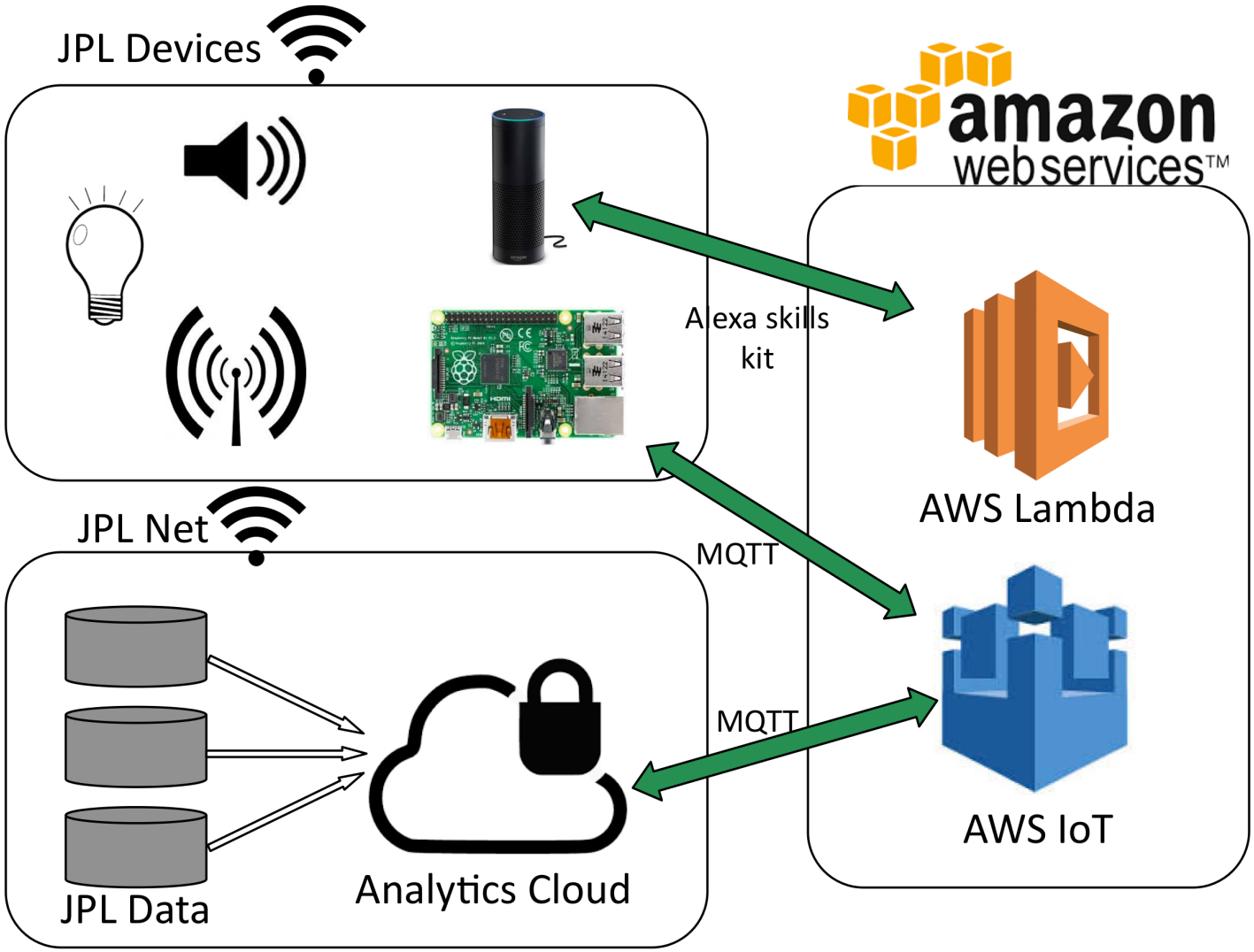
BY KEVIN LISOTA on November 29, 2016 at 10:35 pm

Post a Comment | Share 172 | Tweet | Share 52 | Reddit | Email

Buy tickets to the GeekWire Bash here!

# Device Network

## Addressing Security Concerns




NASA


NASA